A novel Bayesian respiratory motion model to estimate and resolve uncertainty in image-guided cardiac interventions

Devis Peressutti[a,*], Graeme P. Penney[a], R. James Housden[a], Christoph Kolbitsch[a], Alberto Gomez[a], Erik-Jan Rijkhorst[b], Dean C. Barratt[c], Kawal S. Rhode[a], Andrew P. King[a]

[a] *Division of Imaging Sciences and Biomedical Engineering, King's College London, King's Health Partners, St. Thomas' Hospital, London, SE1 7EH, United Kingdom.*
[b] *Department of Physics and Medical Technology, VU University Medical Center, Amsterdam, The Netherlands.*
[c] *UCL Centre for Medical Image Computing, Department of Medical Physics & Bioengineering, University College London, London, UK, WC1E 6BT.*

Abstract

In image-guided cardiac interventions, respiratory motion causes misalignments between the pre-procedure roadmap of the heart used for guidance and the intra-procedure position of the heart, reducing the accuracy of the guidance information and leading to potentially dangerous consequences. We propose a novel technique for motion-correcting the pre-procedural information that combines a probabilistic MRI-derived affine motion model with intra-procedure real-time 3D echocardiography (echo) images in a Bayesian framework. The probabilistic model incorporates a measure of confidence in its motion estimates which enables resolution of the potentially conflicting information supplied by the model and the echo data. Unlike models proposed so far, our method allows the final motion estimate to deviate from the model-produced estimate according to the information provided by the echo images, so adapting to the complex variability of respiratory motion. The proposed method is evaluated using gold-standard MRI-derived motion fields and simulated 3D echo data for 9 volunteers and real 3D live echo images for 4 volunteers. The Bayesian method is compared to 5 other motion estimation techniques and results show mean/max improvements in estima-

*Corresponding author. devis.peressutti@kcl.ac.uk.

tion accuracy of 10.6%/18.9% for simulated echo images and 20.8%/41.5% for real 3D live echo data, over the best comparative estimation method.

Keywords: Respiratory motion, modelling, Bayesian inference, MRI, US

1. Introduction

Image-guided interventions typically involve the use of pre-procedure images to help surgeons perform procedures more quickly, safely and effectively (Perrin et al., 2009; Cleary and Peters, 2010). In cardiac applications, the image data are used to identify the position of target tissues or clinically relevant structures in the heart. The success of the guidance depends on the accurate alignment of the pre-procedure imaging data and the underlying real anatomy that it represents. High spatial and temporal resolution is desirable. One possible approach is to combine high spatial resolution images acquired in the pre-procedural phase, such as Magnetic Resonance Imaging (MRI) (De Buck et al., 2005; Rhode et al., 2005; Yu et al., 2005) or Computed Tomography (CT) (Sra et al., 2007; Knecht et al., 2008; Ector et al., 2008), with high temporal resolution X-ray fluoroscopy images acquired during the procedure. Interest in the use of echocardiography (echo) imaging as an intra-procedure modality has grown over the last few years (Grau et al., 2007; Linte et al., 2008; Wein et al., 2009; Ma et al., 2010; Noble et al., 2011; Gao et al., 2012) due to its high temporal resolution, non-invasive nature and low cost. However, echo images are affected by various artefacts and echo image quality strongly depends on the sonographer's experience and the subject's anatomy.

A significant limitation of image-guided interventions applied to organs in the chest and abdomen is respiratory motion. This causes misalignments between the pre-procedure images used for guidance and the underlying real anatomy, making the guidance information misleading and potentially dangerous (Hawkes et al., 2005). Cardiac respiratory motion contains significant variation, and the motion of the heart differs from exhalation to inhalation (*intra-cycle variability*), and from cycle to cycle (*inter-cycle variability*) (Keall et al., 2006; Blackall et al., 2006). Successful compensation for respiratory motion is a very challenging task but is necessary to fulfil the accuracy requirements of many clinical applications (Linte et al., 2010).

Motion models have been proposed as a possible solution. Their aim is to estimate and model the motion of the organ in order to update the

pre-procedure image data to compensate for the motion, thus maintaining the accuracy of the guidance information. The typical approach is to model the respiratory motion as a direct function of some physically measurable input signal(s), or respiratory surrogate(s). During the intervention, only the surrogate data need be acquired, and the motion model produces a motion estimate based on this data (Manke et al., 2003; Shechter et al., 2005; King et al., 2009; Klinder et al., 2010). Motion models which use this approach are often referred to as surrogate-driven motion models.

In an attempt to improve the accuracy of motion models, there has been growing interest in recent years in combining motion models with intra-procedure imaging data. Such models, also known as image-driven models, do not require physically measurable surrogates. Rather, the respiratory motion is parameterised as a function of some internal variables, such as the position in the respiratory cycle (Blackall et al., 2005; King et al., 2010) or the weights of a statistical model (Schneider et al., 2010; King et al., 2012). During the procedure, imaging data are used to drive the model and internal variables are optimised so that the model output matches the imaging data.

To the authors' knowledge, all motion models proposed to date have a significant limitation in that, for a given value of the surrogate data or internal variables, only a single predetermined motion estimate can be produced. In probabilistic terms, the model outputs the same motion estimate with a 100% confidence in estimation. In reality, the significant intra- and inter-cycle variation of respiratory motion means that such a restrictive model may introduce errors into the estimation process.

Over the last few years, the concept of uncertainty has been investigated and employed in several fields of medical image processing. For instance, confidence measures have played a significant role in optical flow measurements to identify unreliable flow vectors and recover corrupted optical flow regions (Kondermann et al., 2008). Taron et al. (2005) presented a method for obtaining an uncertainty measure from the registration process of different shapes. Similarly, Blanc et al. (2009) proposed a technique for estimating confidence regions around statistical shape models from partial observations. In this paper, we combine the concepts of surrogate-driven and image-driven motion models with the concept of uncertainty to overcome the limitations of previous motion models.

We propose a novel technique to estimate cardiac respiratory motion based on a combination of an MRI-derived pre-procedure motion model and intra-procedure real-time echo images. The main novelty of the proposed

approach lies in its probabilistic formulation, allowing motion estimates to be made with corresponding confidence measures. The use of a Bayesian approach permits combination of the potentially conflicting information provided by the model and the echo data. This approach can be seen as a hybrid surrogate-driven and image-driven model since it employs a measurable surrogate signal for the model formation but uses both the surrogate and imaging data to drive the model. The MRI-derived model estimates the motion and its uncertainty which is then resolved using the imaging data.

Preliminary work has been described in Peressutti et al. (2012) featuring evaluation on simulated echo images. In this paper, the Bayesian model is evaluated on more simulated echo datasets and on real echo imaging data. Refinements to the probabilistic model are also presented.

The paper is structured as follows: the method is described in Section 2, where the prior and likelihood probability distribution functions are defined. Section 3 describes the experiments carried out to evaluate the technique on simulated and real echo images. Results of the experiments are reported in Section 4 while discussion of the results, possible applications of the technique and conclusions are presented in Section 5.

2. Method

An overview of our approach for respiratory motion estimation is presented in Figure 1.

Before the procedure, ECG-gated free-breathing dynamic 3D MRI images are acquired (see Section 3.1). A respiratory surrogate signal is acquired during the dynamic MRI scan. Affine registration of the dynamic images is used to estimate the motion of the heart due to respiration. Each of the affine parameters is modelled as a function of the surrogate signal and an associated uncertainty function is estimated. This serves as the prior probability function (see Section 2.1.1). During the procedure, live 3D echo imaging data are used to form the likelihood term (see Section 2.1.2). A measure of the echo image content is used to determine the optimal combination of prior and likelihood (see Section 2.2). The final motion estimate is obtained in a *Maximum a Posteriori* (MAP) manner (see Section 2.1).

2.1. Bayesian motion estimation

A Bayesian framework is used to combine prior knowledge of the subject-specific respiratory motion, in the form of a MRI-derived affine motion model,

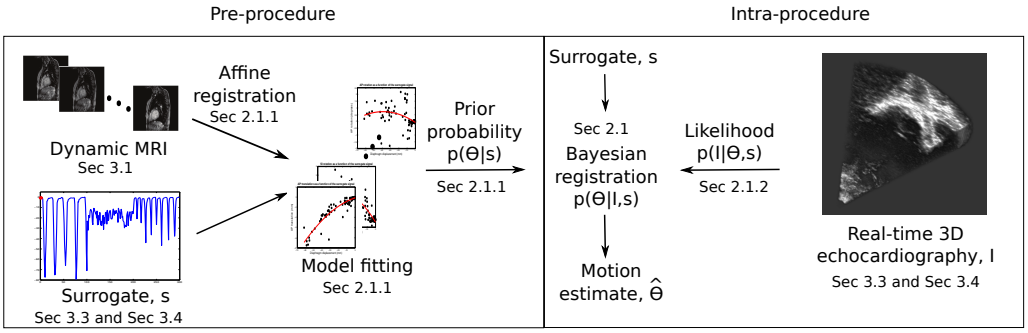


Figure 1: Schematic representation of the Bayesian motion estimation. Before the procedure, an affine motion model is derived from dynamic MRI images and a respiratory surrogate signal, forming the prior probability function. During the procedure, real-time 3D echo images are acquired, forming the likelihood probability function. The respiratory surrogate is acquired as well, providing a starting estimate for the Bayesian optimisation. The posterior probability is then maximised, obtaining the final motion estimate.

with real-time echo images.

The posterior probability represents the state of certainty about a given phenomenon. The phenomenon here is the affine transformation $\boldsymbol{\theta}$ that describes the respiratory motion of the heart, where $\boldsymbol{\theta}$ is a vector of 12 affine motion parameters. Denoting by s the respiratory surrogate signal and by I the echo imaging data, Bayes' law states that the posterior probability function is proportional to the product of the likelihood and prior probability functions

$$p(\boldsymbol{\theta}|I,s) = \frac{p(I|\boldsymbol{\theta},s)\cdot p(\boldsymbol{\theta}|s)}{p(I|s)}. \tag{1}$$

The prior probability function $p(\boldsymbol{\theta}|s)$ represents the degree of certainty about the respiratory motion $\boldsymbol{\theta}$ given the surrogate s, before the echo images are considered. The prior is computed from the affine motion model as a function of the surrogate (see Section 2.1.1). The likelihood $p(I|\boldsymbol{\theta},s)$ represents the probability of the echo image I given the affine transformation $\boldsymbol{\theta}$ and surrogate s. Thus, the likelihood encompasses the information carried by the new data, namely, by the real-time echo imaging data I.

As is common in Bayesian image analysis (Geman and McClure, 1985; Hanson, 1993), we adopt the simplifying assumption that the normalising

factor $p(I|s)$ has a uniform distribution. The desired affine motion estimate $\hat{\boldsymbol{\theta}}$ is obtained in a MAP manner,

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta},s}\{p(\boldsymbol{\theta}|I,s)\}. \tag{2}$$

The parameters $\boldsymbol{\theta}$ and the surrogate s are optimised to maximise $p(\boldsymbol{\theta}|I,s)$ given the echo image I. The prior and the likelihood terms are now described in the following sections.

2.1.1. Prior probability

The prior probability is formed from the MRI-derived motion model. To build such a model, affine intensity-based registrations between a reference end-exhale dynamic 3D MRI image and the remaining dynamic 3D MRI images (see Section 3.1) are performed. To constrain the registration to the heart only, an ellipsoidal mask covering the four chambers and the major cardiac vessels was employed. This registration process results in 12 affine parameters for each dynamic image. 2^{nd} order polynomial curves are fitted in a least-squares sense to each affine parameter as a function of the respiratory surrogate signal (King et al., 2009). Given a value s of the surrogate acquired in the intra-procedure setting, the MRI-derived motion model outputs a twelve parameter affine transformation, denoted by $\tilde{\boldsymbol{\theta}}(s)$.

To incorporate a measure of uncertainty into the motion estimates made by the motion model, Gaussian distributions are fitted to the variations of each of the 12 affine parameters away from the model estimate $\tilde{\boldsymbol{\theta}}(s)$. An example of model fitting and a Gaussian function is shown in Figure 2 for the 1^{st} affine parameter, i.e. anterior-posterior (AP) translation. The peak of the Gaussian corresponds to the fitted value $\tilde{\theta}_i(s)$ while the variance $\sigma^2_{p_i}$ is a function of the residuals of the fitting method and the surrogate s.

As a simplifying assumption, the 12 affine parameters are considered to be statistically independent, so the prior probability is given by the product of the 12 Gaussian distributions,

$$p(\boldsymbol{\theta}|s) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2_{p_i}(s)}} e^{-\frac{\left(\theta_i - \tilde{\theta}_i(s)\right)^2}{2\sigma^2_{p_i}(s)}}, \quad i = 1,\ldots,12. \tag{3}$$

where $\sigma^2_{p_i}(s)$ is the variance function for the i^{th} parameter. We now describe in more detail how $\sigma^2_{p_i}(s)$ is determined.

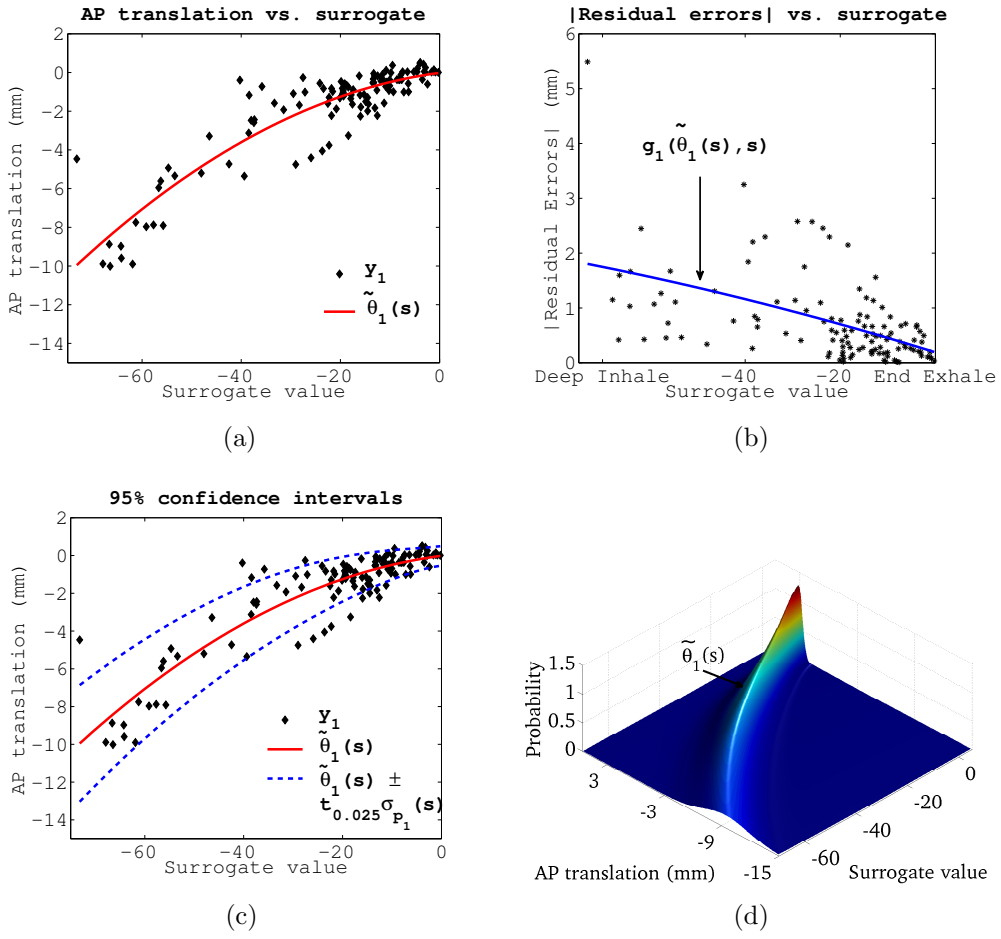


Figure 2: Example of Gaussian distribution function for the first affine parameter (AP translation). (a) shows the 2^{nd} order polynomial function $\tilde{\theta}_1(s)$ that models the observations y_1 as a function of the surrogate s. As shown in (b), $g_1(\tilde{\theta}_1(s), s)$ is used to derive the variance of the Gaussian distribution as a function of the residual error and surrogate s. 95% confidence intervals in prediction (dashed blue lines) (c) and the resulting probability distribution function for the AP translation (d) are shown.

We denote by $\boldsymbol{y_i}$ the observations for the affine parameter i. For instance, Figure 2(a) shows the distribution of $\boldsymbol{y_1}$ as a function of the surrogate s. The linear regression model solves

$$\boldsymbol{y_i} = \boldsymbol{\beta_i} X + \boldsymbol{\epsilon_i}, \;\; i = 1, \ldots, 12 \tag{4}$$

where X is the design matrix having [s^2 s 1] as column vectors and as many columns as the dynamic MRI images, $\boldsymbol{\beta_i}$ are the regression coefficients of the polynomial function defining $\tilde{\theta}_i(s)$ (see Figure 2(a)) and $\boldsymbol{\epsilon_i}$ is the error term for the affine parameter i (Seber and Lee, 2003). A common assumption is that $\boldsymbol{\epsilon_i}$ is normally distributed over the predictor variables (*homoscedasticity*). However, as can be seen in Figure 2(a) and 2(b), the residual distribution is not constant over the surrogate signal s, suggesting a correlation between the variance of the error $\boldsymbol{\epsilon_i}$ and the surrogate s (*heteroscedasticity*). The heteroscedastic behaviour of the affine parameters is due to respiratory motion variability, confirming that respiratory motion at inhale is less repeatable than at end-exhale (see Figure 2(b)). Therefore, a heteroscedastic variance function $\sigma^2_{p_i}(s)$ is proposed

$$\sigma^2_{p_i}(s) = \boldsymbol{x_0}^T S_i \boldsymbol{x_0} + g_i^2(\tilde{\theta}_i(s), s), \;\; i = 1, \ldots, 12. \tag{5}$$

$\sigma^2_{p_i}(s)$ is the variance of the error in prediction $y_0 = \hat{\boldsymbol{\beta}}_{\boldsymbol{i}} \boldsymbol{x_0}$, where the first term of Eq. 5 represents the error due to the data used to compute $\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}$ while the second term represents the variance of the error term $\boldsymbol{\epsilon_i}$. For large sample sizes, the first term is negligible (Davidian and Carroll, 1987). S_i is the covariance matrix of the regression coefficients estimate $\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}$

$$S_i = \hat{\sigma}_i^2 (X^T X)^{-1}, \;\; i = 1, \ldots, 12 \tag{6}$$

where $\hat{\sigma}_i^2$ is the mean squared error for the affine parameter i, X is the design matrix, $\boldsymbol{x_0}$ is the predictor vector for the new observation s_0 ($\boldsymbol{x_0} = [\mathrm{s}_0^2 \; \mathrm{s}_0 \; 1]$) while $g_i^2(\tilde{\theta}_i(s), s)$ describes the variance of the model error $\boldsymbol{\epsilon_i}$ as a function of s. As proposed by Davidian and Carroll (1987), $g_i(\tilde{\theta}_i(s), s)$ is a function fitted to the absolute values of the residuals for the parameter i (see Figure 2(b)). 2^{nd} order polynomial functions and least-squares fitting were again employed. This heteroscedastic model allows us to generate a probability density function with a variance that depends on the surrogate value, as shown in Figure 2(d).

2.1.2. Likelihood

The likelihood function $p(I|\boldsymbol{\theta}, s)$ estimates the probability that the acquired image I would be generated by a given affine motion $\boldsymbol{\theta}$. To compute the likelihood, an end-exhale reference echo image I_{ref} was selected for each dataset (see Section 3). A Gaussian distribution based on the similarity measure between the new echo image I and the affine transformation of the reference echo image $A(I_{ref}, \boldsymbol{\theta})$ was adopted,

$$p(I|\boldsymbol{\theta}, s) = \frac{1}{\sqrt{2\pi\sigma_l^2}} e^{-\frac{\left(NCC(A(I_{ref},\boldsymbol{\theta}),I)-1\right)^2}{2\sigma_l^2}} \tag{7}$$

where $NCC(A(I_{ref}, \boldsymbol{\theta}), I)$ is the value of the normalized cross correlation similarity measure between the affine warping of the reference image $A(I_{ref}, \boldsymbol{\theta})$ and the new echo image I, and σ_l^2 is the likelihood variance. The more similar I and $A(I_{ref}, \boldsymbol{\theta})$ are ($NCC(A(I_{ref}, \boldsymbol{\theta}), I) \to 1$), the higher the value of the likelihood $p(I|\boldsymbol{\theta}, s)$. An automatic method for determining the optimal value of σ_l^2 is now described.

2.2. Optimisation of σ_l^2

The values of σ_l^2 and $\sigma_{p_i}^2$ affect the final motion estimate by changing the peak of the posterior probability function. For instance, using a non-informative prior, i.e. a uniform distribution function ($\sigma_{p_i}^2 \to \infty$), would lead to a motion estimate completely driven by the likelihood function, whereas, on the other hand, greater values of σ_l^2 would lead to a motion estimate tending to the model estimate $\tilde{\boldsymbol{\theta}}(s)$.

As detailed in Section 2.1.1, $\sigma_{p_i}^2$ was determined by the model data, i.e. the residuals of the modelling process. Therefore, the only free parameter that needs to be chosen in our technique is σ_l^2.

A reasonable assumption is that the value of σ_l^2 should reflect the quality of the information provided by the new echo imaging data, i.e. the quality of the acquired echo images, I. Ideally, if the echo imaging data were good enough to fully characterise the position of the heart, the prior knowledge from the model would not be necessary.

Quantification of image quality is a controversial and challenging topic, especially for echo imaging, due to its characteristic speckle pattern and many other artefacts such as shadowing, spatial distortion and multiple reflections (Kremkau and Taylor, 1986; Thijssen, 2003; Noble and Boukerroui, 2006). In this work, the relationship between σ_l^2 and a differential signal-to-noise ratio

(SNR_d) of the echo images is investigated, resulting in a simple empirical expression that can be used to determine the optimal value of σ_l^2 before the Bayesian optimisation is performed.

Since the left ventricle (LV) was well imaged by both the dynamic MRI and echo images, the following differential SNR_d measure was considered

$$SNR_d = \frac{\mu_{myo} - \mu_{blood}}{\sigma_{noise}}, \quad \sigma_{noise} = \frac{\sigma_{myo} + \sigma_{blood}}{2} \tag{8}$$

where μ_{myo} and σ_{myo} are the mean and standard deviation of the LV myocardium voxel values while μ_{blood} and σ_{blood} are the mean and standard deviation of the blood voxel values in the LV cavity. Higher values of SNR_d mean that LV myocardium voxels have more homogeneous values that differ markedly from voxels of the blood, therefore the echo image is highly informative. On the other hand, low values of SNR_d indicate poor contrast images with more heterogeneous values of the LV myocardium voxels.

For each echo image, the SNR_d is computed and the empirical model used to determine a value for σ_l^2. This value of σ_l^2 is used in the optimisation of the posterior probability. Section 3.5 provides further details on how the determination of the optimal σ_l^2 was carried out for the simulated and real echo images.

3. Experiments

The Bayesian technique proposed was tested on simulated echo images (see Section 3.3) and on real echo imaging data (see Section 3.4). The use of simulated echo images, with realistic MRI-derived motion fields, enabled us to perform a thorough accuracy/robustness evaluation with known, gold-standard, motion fields. The use of real echo images enabled evaluation on data similar to those which would be acquired in a clinical environment, but without known, gold-standard, motion fields.

In both cases, the prior probability term was formed from the MRI-derived affine motion model. Details of the MRI sequences are provided in Section 3.1.

3.1. MRI acquisition

The MRI images were acquired using a 1.5T Philips Achieva MRI scanner. The dynamic 3D MRI sequence was used for forming the motion model :

Dynamic 3-D : 3-D TFEPI, ECG-triggered and gated at late diastole, typically 20 slices, $TR = 10ms$, $TE = 4.9ms$, flip angle = 20°, acquired voxel size 2.7 x 3.6 x $8.0mm^3$, reconstructed voxel size 2.22 x 2.22 x $4.0mm^3$, TFE factor 26, EPI factor 13, TFE acquisition time $267.9ms$.

The dynamic 3D sequence was ECG-triggered and gated, so one volume was acquired for each heart beat. The images therefore represented the motion of the heart due to respiration only. Compared to the dynamics of respiration, the short acquisition time ($267.9ms$) allows acquisition of near motion-free 3D images. 2-4 dynamic volumes were typically acquired for each breathing cycle. During the dynamic scan, the subject was asked to breathe in three different ways: normal, fast and deep breathing. This allowed respiratory variability to be captured and modelled. The overall acquisition of the dynamic MRI images lasted approximately 2-5 min.

For this study, 13 volunteer datasets were processed, 9 datasets for simulated echo images (Vol. A-I) and 4 datasets for real echo data (Vol. J-M). For volunteers A-D, the dynamic sequence acquired 300 images (100 images for each different breathing pattern) while for the remaining volunteers E-M, the dynamic sequence acquired 120 images (40 images for each breathing pattern). In the latter case, the number of images was sufficient to represent an adequate number of respiratory cycles (at least 4) for each breathing pattern. The short acquisition time for each dynamic image ($267.9ms$) put restrictions on the coverage of the dynamic scans, therefore the field of view covered most, but not all, of the four chambers of the heart.

In addition, for each of the 13 volunteer datasets considered, a high resolution 3D MRI volume was acquired :

High resolution 3-D : 3-D balanced TFE, cardiac gated at late diastole, respiratory gated at end-exhale, $5mm$ navigator window, typically 120 sagittal slices, $TR = 4.4ms$, $TE = 2.2ms$, flip angle= 90°, acquired voxel size 2.19 x 2.19 x $2.74mm^3$, reconstructed voxel size 1.37 x 1.37 x $1.37mm^3$, the acquisition window was optimised for each volunteer and was on average $100ms$, scan time approximately 5 minutes.

The high resolution MRI image is a standard pre-procedure acquisition in many clinical protocols and provides high spatial resolution information about the anatomy and pathology of the heart. For the simulated echo experiments, the high resolution MRI volume was segmented and used in the echo image simulation (see Section 3.3), while for the real echo experiments

it was used for evaluation purposes (see Section 3.4). Also, the segmentation of the LV from the high resolution MRI image was employed for computing the SNR_d of the echo images (see Section 3.5).

3.2. Comparison of estimation techniques

To evaluate the performance of the proposed method, for both simulated (see Section 3.3) and real echo experiments (see Section 3.4), several motion estimation approaches were compared:

no estimate: the reference end-exhale position is used as motion estimate for any inhale position;

model-only: a surrogate-driven model estimate (King et al., 2009). Given the surrogate value s, the motion estimate is $\tilde{\boldsymbol{\theta}}(s)$. Details of precisely which surrogate signals were used in our experiments are provided in Sections 3.3 and 3.4;

image-only: as in King et al. (2010), the estimate is given by the registration of the echo images only, without the constraints of a motion model. This approach corresponds to the Bayesian estimate using a non-informative prior. However, unlike a purely image-only registration, $\tilde{\boldsymbol{\theta}}(s)$ was provided as a starting estimate for the optimisation algorithm since the small field of view and the even smaller overlapping region of $A(I_{ref}, \boldsymbol{\theta})$ and I made the optimisation unstable. A hill climbing optimisation algorithm was used;

PCA-based model: similar to Schneider et al. (2010) and King et al. (2012), the set of affine parameters $y_i,\ i = 1, \ldots, 12$ is assumed to lie on a linear sub-manifold having dimensionality $d \leq D,\ D = 12$. Linear PCA is applied to find the first d eigenvalues $\lambda_1 \geq \ldots \geq \lambda_d$ and corresponding eigenvectors $\mathbf{v_1}, \ldots \mathbf{v_d}$ that capture at least 95% of the overall variance $\sum_{j=1}^{D} \lambda_j$. Denoting by $\mathcal{Y} = [\boldsymbol{y_1}; \ldots; \boldsymbol{y_D}]$ the $D \times H$ matrix where $D = 12$ and $H = 300-1$ for Vol. A-D and $H = 120-1$ for Vol. E-M, singular value decomposition is applied to the covariance matrix $\Sigma = \mathcal{Y}_c \mathcal{Y}_c^T$, where $\mathcal{Y}_c = \mathcal{Y} - \bar{\boldsymbol{\mathcal{Y}}}$ is centred with respect to the D-component mean vector $\bar{\boldsymbol{\mathcal{Y}}} = \frac{1}{H} \sum_{h=1}^{H} \mathcal{Y}$. Given the difference in units of the 12 affine parameters, the values of each affine parameter are normalised with respect to their standard deviation before PCA is applied. The PCA-based model estimate $\hat{\boldsymbol{\theta}}_{\boldsymbol{PCA}}$ is obtained by optimising the d-component

vector $\boldsymbol{\alpha}$ of weights that maximises

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{PCA}}(\boldsymbol{\alpha}) = \arg\max_{\boldsymbol{\alpha}}\{NCC(A(I_{ref}, \boldsymbol{\theta}_{\boldsymbol{PCA}}(\boldsymbol{\alpha})), I)\}, \tag{9}$$

where $\boldsymbol{\theta}_{\boldsymbol{PCA}}(\boldsymbol{\alpha}) = \bar{\boldsymbol{\mathcal{Y}}} + \sum_{j=1}^{d} \alpha_j \boldsymbol{v}_j$. A hill climbing optimisation algorithm was used;

simple model and image: using a method similar to that described in Blackall et al. (2005) and King et al. (2010), the estimate is obtained by maximising $NCC(A(I_{ref}, \boldsymbol{\theta}), I)$ but constraining $\boldsymbol{\theta}$ to lie on the model, that is

$$\hat{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}(\hat{s}), \;\; \hat{s} = \arg\max_{s}\{NCC(A(I_{ref}, \tilde{\boldsymbol{\theta}}(s)), I)\}. \tag{10}$$

This approach corresponds to an image-driven model, without providing either a measure of uncertainty nor deviating from $\tilde{\boldsymbol{\theta}}(s)$. A hill climbing optimisation algorithm was used;

Bayesian: our proposed motion estimate (see Equation 2). As described in Section 3.5, the optimal value of the likelihood variance was determined by means of Eq. (12) for all simulated and real echo images. In the optimisation, the model-only estimate $\tilde{\boldsymbol{\theta}}(s)$ was used as a starting estimate. A hill climbing algorithm was used to estimate $\hat{\boldsymbol{\theta}}$.

3.3. Evaluation using real MRI, simulated echo data

The Bayesian motion model was evaluated on simulated echo images derived from the high resolution MRI image and the dynamic MRI images employed to build the model, as described in Peressutti et al. (2012). The evaluation presented here adds 5 extra volunteer datasets, and more realistic fields of view (FOV) and acoustic windows for the simulated echo images.

A block diagram illustrating the evaluation using simulated echo imaging with realistic, gold-standard, motion fields is shown in Figure 3.

The superior-inferior (SI) displacement of the dome of the left hemidiaphragm was chosen as the respiratory surrogate signal and automatically extracted from all dynamic MRI images, simulating an MRI navigator echo (Savill et al., 2011). As described in King et al. (2009), the *SI* diaphragm displacement as imaged by X-ray fluoroscopy images can also be employed as the intra-procedure surrogate signal.

The dynamic MRI image having the highest SI displacement was selected as the end-exhale reference image for registration. In order to have

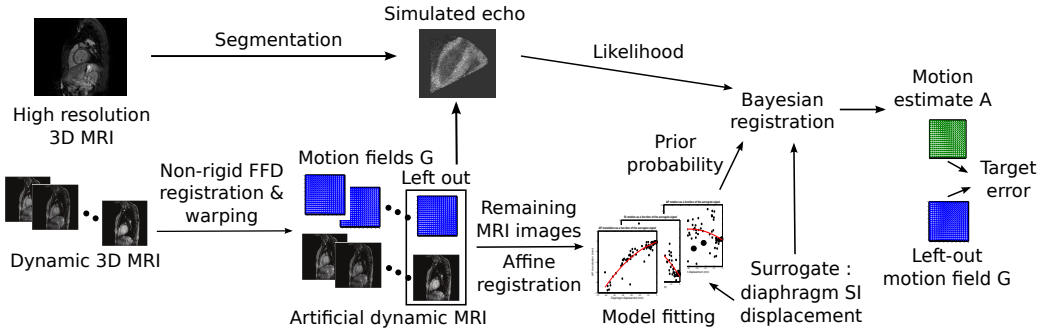


Figure 3: Block diagram of the echo simulation and leave-one-out validation framework. Real dynamic MRI images are non-rigidly registered to a dynamic end-exhale reference image, and warped using the resulting motion fields to create artificial, but realistic, dynamic MRI images with known gold standard motion fields. Each artificial MRI image is left out in turn. The non-rigid motion field of the left out dynamic is combined with a segmented high resolution MRI image to produce a simulated 3D echo image. By using an intensity-based affine registration algorithm, the remaining artificial MRI images are used to form the motion model. To assess the accuracy, the affine motion fields A are estimated by the model and then compared with the left-out ground truth non-rigid motion fields G.

a gold-standard set of deformation fields to simulate echo images and validate the motion model, a free-form deformation (FFD) registration algorithm (Buerger et al., 2011) was used to register each dynamic MRI image to the end-exhale reference image. The non-rigid registration resulted in a set of ground truth deformation fields G and the corresponding 'artificial' set of warped 3D MRI images was used to build the motion model and form the prior probability (see Section 2.1.1). Thus, in this evaluation, the intensity-based affine registration was applied to the 'artificial' set of warped 3D MRI images instead of the original dynamic sequence. This approach was employed to obtain a gold standard set of motion fields where all deformations were known and used for accuracy/robustness evaluation.

3D echo images were simulated using the ultrasound propagation model proposed in Rijkhorst et al. (2010). This method requires maps of the 3D acoustic impedance Z and absorption coefficients α corresponding to different tissues. The high resolution 3D MRI image was manually segmented to differentiate blood pool, myocardium, liver and lungs. Due to poor contrast,

	blood	heart	muscle	lungs	liver
Z (10^6 kg s^{-1} m^{-2})	1.61	1.71	1.70	0.50	1.65
α (dB cm^{-1})	0.18	1.10	1.00	12.00	0.94

Table 1: Values of the acoustic impedance Z and absorption coefficient α used for simulating B-mode echo images.

it was not possible to segment the ribs out of the high resolution volume.

The values of the acoustic impedance and absorption coefficients were assigned according to Rijkhorst et al. (2010) and Baun (2010) and are reported in Table 1. The echo simulation method uses a Lambertian reflection model simulating reflection, absorption and transmission ratios along each ray direction. The effects of finite beam width in the elevational direction and of multiple active transducer elements are also simulated. To simulate realistic speckle noise, two different textures, one for the myocardium and one for the other tissues, were used. These textures were obtained using a texture quilting technique (Efros and Freeman, 2001) applied to real 3D echo patches of the myocardium and the blood pool, acquired from a volunteer. To obtain a SNR_d comparable to real echo images, Gaussian noise was added to the simulated echo images based on measurements of real echo images, as described in Section 3.5.

The high resolution MRI volume was rigidly transformed to the coordinate system of the dynamic reference MRI image and the segmentation was then warped to each dynamic inhale position using the FFD motion fields from the non-rigid registrations. The acoustic impedance and absorption maps were assigned to the warped segmentations and 3D echo images were simulated. A trans-apical acoustic window was adopted with a FOV that is typical of images acquired in real-time from a modern echo machine (see Section 3.4). The US simulation of the dynamic end-exhale reference MRI image was employed as the echo reference image I_{ref} in the computation of the likelihood term (see Section 2.1.2).

Nine volunteer datasets were processed. A leave-one-out framework was employed to validate the Bayesian approach using simulated echo images. This means that, for each simulated echo image, the corresponding non-rigid motion fields and 'artificial' MRI image were not utilised for building the motion model and forming the prior probability (see Section 2.1.1).

Given the availability of the ground truth non-rigid motion fields G, target

registration errors were used to assess the accuracy of the Bayesian model and the other comparative approaches (see Section 3.2). A set of K landmarks ($\approx$ 35,000) was positioned on a regular Cartesian grid within a manually delimited elliptical mask covering the four chambers and the major vessels of the heart (see Section 2.1.1) in the end-exhale reference image, $\vec{x}_k, 1 \leq k \leq K$. Using the left-out ground-truth motion fields G, the set of landmarks was warped in turn to each of the H dynamic MRI inhale phases, with H being either $H = 300-1$ or $H = 120-1$ (see Section 3.1). Denoting by A the affine transformation estimated by each of the techniques compared, the error in estimation was computed as

$$\|G_h(\vec{x}_k) - A_h(\vec{x}_k)\|, \quad 1 \leq k \leq K, \quad 1 \leq h \leq H. \tag{11}$$

Median and 95^{th} quantiles were adopted as statistics because of the skewness of the error distributions. Results are reported in Section 4.1.

3.4. Evaluation using real MRI, real echo data

To evaluate the Bayesian motion model on real live echo images, experiments were carried out in an XMR catheterisation suite (Rhode et al., 2003, 2005). This system allows automatic determination of the MRI to patient rigid transformation, employing an Optotrak 3020 optical tracking system (Northern Digital Inc.) to track the patient bed. The workflow used is illustrated in Figure 4. While in the case of simulated echo images the SI displacement of the diaphragm was employed as a respiratory surrogate signal (see Section 3.3), for real echo images a respiratory bellows placed on the chest of the subjects was used to form the prior probability $p(\boldsymbol{\theta}|s)$ and to optimise the posterior probability $p(\boldsymbol{\theta}|I, s)$. The respiratory bellows signal can be easily acquired during the procedure and used as the intra-procedure surrogate signal s.

3D live echo images were acquired using an iE33 3-D real-time echocardiography system with a X3-1 3 to 1 MHz broadband matrix array transducer (Philips Healthcare). Light emitting diodes attached to the echo probe were tracked using the Optotrak tracking system, allowing registration of the echo images to the physical space of the XMR suite. The image-to-probe calibration method proposed by Ma et al. (2009) was employed. Following the registration framework described in King et al. (2010), the MRI to echo imaging rigid transformation was computed. Since tracking data, echo imaging data

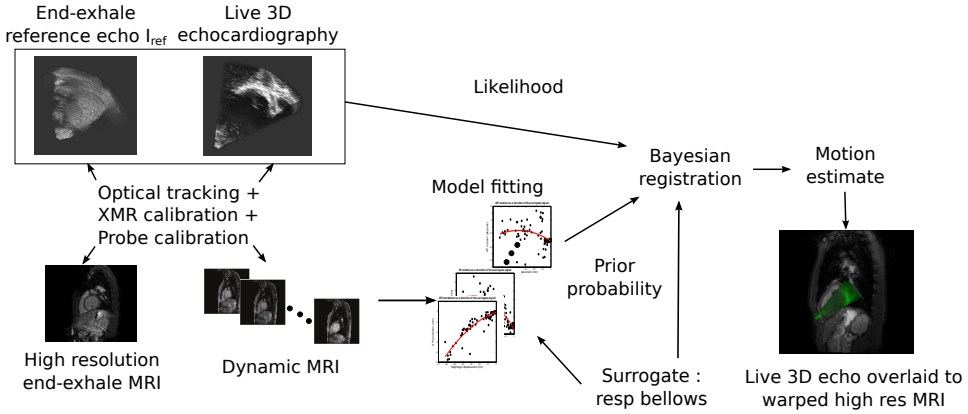
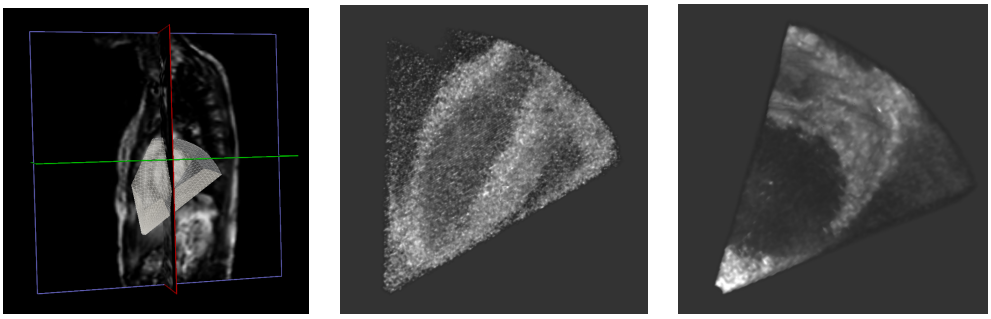


Figure 4: Workflow employed to evaluate the technique on real echo imaging data. The dynamic MRI images were used to build the affine motion model. A respiratory bellows was employed as surrogate signal. The MRI-derived affine motion model formed the prior probability, whereas the full-volume 3D reference and 3D live echo images formed the likelihood term. Optical tracking together with calibration of the XMR catheter laboratory and the echo probe provided the rigid transformation between the MRI and the echo imaging coordinate system. For evaluation purposes, the high resolution MRI image was transformed to the dynamic MRI coordinate system and then warped with the motion estimates produced by the different approaches compared (see Section 3.2). The corresponding live 3D echo was overlaid onto the warped MRI image and misalignment was visually assessed.

and respiratory bellows data were acquired on different machines, synchronisation of the machine clocks and time stamping of all data was performed to ensure temporal correspondence.

The geometric characteristics of the 3D live ultrasound beam, in pyramidal coordinate system, were: inclination angle = 30°, inclination offset = $-1°$, azimuthal angle = 55°, azimuthal offset = -27.5°, radial depth = $140mm$. These same parameters were used to simulate the 3D echo images in the previous Section 3.3 to improve the realism of the simulated images and assess the similarity between simulated and real echo images. An example of beam shape, a simulated echo image and a real echo image are shown in Figure 5. For the echo machine employed, the 3D standard echo images are acquired over four beating cycles of the heart, i.e. they are compounded from 4 smaller volumes, each acquired in a single heart beat. Therefore,

the volume imaged by a standard echo image is four times bigger than the volume imaged by 3D live echo images.



(a) (b) (c)

Figure 5: (a) Example of 3D live ultrasound beam overlaid onto a dynamic MRI image; (b) 2D section of a simulated 3D echo image; (c) reformatted 2D slice through a live 3D echo image.

Four volunteer datasets were acquired. For each dataset, 4 standard echo images were acquired during end-exhale breath hold and 12 sequences of free-breathing live echo images were acquired. Of the 12 free-breathing sequences, 3 sequences were acquired during normal breathing, 3 during fast breathing and 3 during deep breathing. Each sequence lasted approximately 4 seconds and live echo images were streamed at a rate of 14 images per second. However, since the MRI-derived motion model describes the heart's position at end-diastole, the echo images were manually retrospectively gated at end-diastole, resulting in 4 to 6 live echo images per sequence. Modified parasternal acoustic windows were used and the standard end-exhale echo image covering a similar FOV to the live images was employed as the reference image for the computation of the likelihood term (see Section 2.1.2).

By using the MRI to patient rigid transformation, the standard full-volume reference echo images and the live echo images were transformed to the coordinate system of the dynamic MRI images, where the affine motion model was represented. Given the value of the respiratory bellows corresponding to each live echo image, the respiratory position was estimated by each of the different estimation methods (see Section 3.2).

To assess the accuracy of the motion estimates, the high resolution MRI image was first rigidly transformed to the dynamic MRI coordinate system

and then warped to the affine inhale position estimated by each approach. The corresponding live echo image was overlaid onto the warped high resolution MRI image and the maximum misalignment error of corresponding anatomical landmarks (e.g. inter-ventricular septum, LV posterior wall) was visually assessed. As they had higher image contrast, anatomical landmarks on the lateral/posterior LV wall were mainly considered for error assessment, while the position of the septum was used to assess the overall realism of the motion estimate. The GIMIAS open-source visualisation software (Larrabide et al., 2009) was used to estimate the misalignment. Three independent observers were asked to assess 200 pairs of images (i.e. 50 pairs of images for each volunteer, 10 for each of the estimation techniques compared (see Section 3.2)) randomly selected and measure the maximum misalignment error. Observers were also asked to report any obviously unrealistic motion estimations. The evaluation assessment was blinded, that is the observers were presented with the image pairs without knowing which technique they were produced with. An example of misalignment assessment is shown in Figure 6. Results are reported in Section 4.

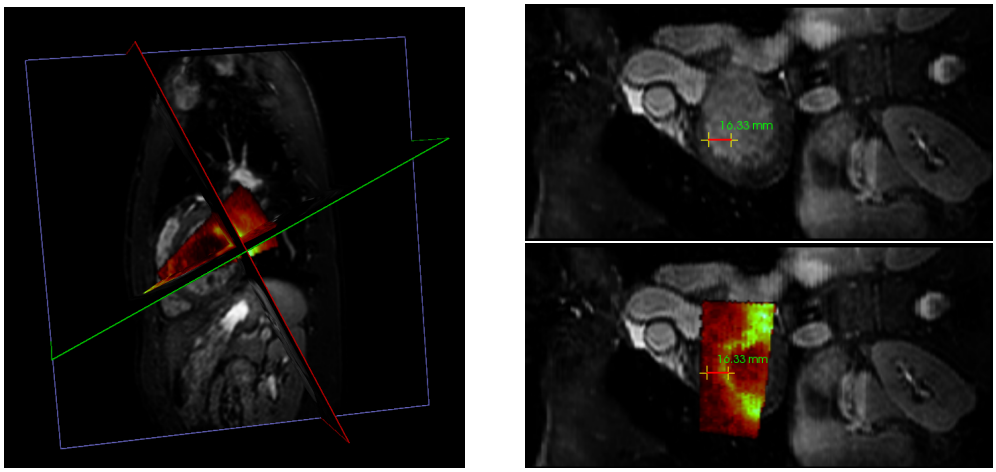


Figure 6: Example of evaluation assessment using GIMIAS. The 3D live echo image is overlaid onto the transformed high resolution MRI image and the maximum misalignment error (shown in green) is measured. 3D plane orientation is shown on the left. On the right, error assessment shown on the transformed high resolution MRI image (top) and on the 3D live echo (bottom) shown in a modified coronal view.

3.5. Determination of the optimal σ_l^2

To automatically determine the optimal value of σ_l^2, we employ a simple linear model based on the SNR_d of the echo images. This means that each echo image will potentially have a different value of σ_l^2 used in the optimisation, depending on the quality of the image. We now describe the formation of this simple model based on the simulated and real echo images.

We first need to compute the SNR_d for each echo image. For simulated echo images (see Section 3.3), LV myocardium and blood were known from the segmentation of the high resolution MRI image used for echo simulation (see Section 3.3). For each dataset, two inhale echo images were selected and Gaussian noise having different values of variance was added in the simulation process to vary the SNR_d of the simulated images. Twelve evenly distributed values of σ_l^2 were selected and the Bayesian estimation for each image and for each value of σ_l^2 was computed. For each image, the value of σ_l^2 producing the lowest target registration error (see Section 3.3) was determined.

For real echo images (see Section 3.4), LV myocardium and blood were segmented from the high resolution MRI image and transformed to the dynamic MRI coordinate system. Given the availability of the image-only registration between the full-volume reference echo and the live echo images (see Section 3.2), the LV segmentation was warped to each live echo image using the image-only registration and the SNR_d was computed for each live echo image. For each of the four volunteers, approximately 15 images were selected and, as for the simulated echo, 12 evenly distributed values of σ_l^2 were selected and the Bayesian estimation computed. The value of σ_l^2 generating the lowest misalignment error (see Section 3.4) was selected. Figure 7 shows the distribution of σ_l^2 producing the best results for simulated echo images and real echo images as a function of the SNR_d. As can be seen, the overlap of the two distributions is high, confirming the similarity between the simulated and real echo images. A linear function was employed to describe the relationship between σ_l^2 and SNR_d

$$\sigma_l^2 = a_1 \cdot SNR_d + a_2. \tag{12}$$

A linear function represented the best trade-off between the quality of fit and the simplicity of representation. Furthermore, the possible range of SNR_d values for real echo images ($0 < SNR_d < 1.5$) and simulated echo images ($1 < SNR_d < 1.3$) ensured $\sigma_l^2 > 0$.

To separate data used for parameter estimation and data used for evaluation, a_1 and a_2 were computed using a leave-one-out test, which means that,

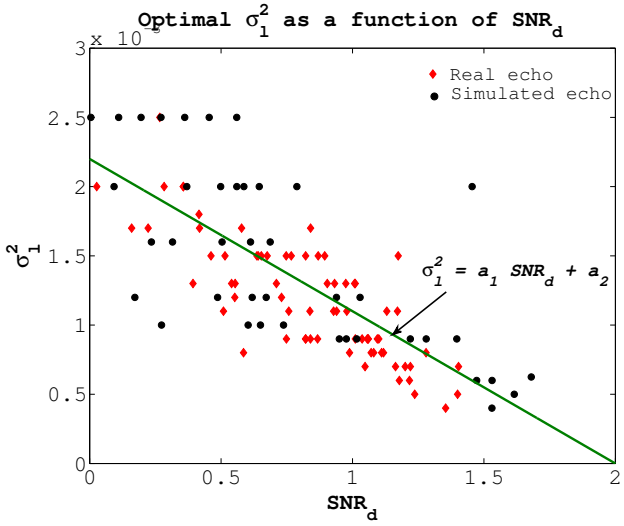


Figure 7: Distribution of the best values of σ_l^2 as a function of SNR_d. Black circles refer to simulated echo images while red diamonds refer to real echo images. A good overlap between simulated and real echo images can be noticed. A linear function is fitted to σ_l^2, defining the optimal value of the likelihood's variance as a function of the echo image's SNR_d.

for any given dataset, the corresponding a_1 and a_2 were fitted without considering the best σ_l^2 derived from the same dataset. A least-squares fitting was used and typical values were $a_1 = -1.1 \cdot 10^{-3}$ and $a_2 = 2.2 \cdot 10^{-3}$, with a correlation coefficient of 0.75 (see Figure 7).

When applying the Bayesian model, for each dataset of simulated and real echo images, the SNR_d of each image was determined as described above and, given a_1 and a_2, the optimal σ_l^2 was computed image-wise (see Eq. 12) and used for the likelihood term of the Bayesian estimation method.

4. Results

Results achieved for simulated and real echo images are presented in Section 4.1 and 4.2, respectively. The median values rather than the mean values are employed as statistics since the error distributions were skewed and non-Gaussian. 95^{th} quantiles represent the worst case scenario and are therefore paramount in clinical applications.

4.1. Simulated echo data

The median and 95^{th} quantile of the target error for each of the estimation techniques compared are summarized in Figure 8 and reported in Table 2.

For each of the 9 datasets considered, the accuracy error achieved by the Bayesian method was lower than any other comparative approach. To summarise the performance of the Bayesian estimation, we also computed the mean/max improvements of median/95^{th} quantile. The mean is the average improvement of the median error over all 9 datasets. The mean/max improvements achieved were 10.6%/18.9% over the best comparative technique. The best comparative technique was defined to be the technique having the lowest 95^{th} quantile, which was, for 8 volunteers out of 9, the combination of a motion model and image data. Mann-Whitney U-tests were performed between the Bayesian error distribution and the best comparative technique. For all volunteers, statistical significance was found with $p < 0.05$.

4.2. Real echo data

Table 3 reports the median and 95^{th} quantile of the maximum misalignment error measured by the three independent observers. Given the lower number of samples for real data compared to simulated data, rather than median/95^{th} quantile, all misalignment errors for Vol. J-M are shown in Figure 9.

As for simulated echo images, the Bayesian method outperformed all the other estimation methods, with mean/max improvements in estimation accuracy of 20.8%/41.5% over the best comparative approach. The best comparative technique was again defined as the technique having the lowest 95^{th} quantile. The mean value is the average improvement of the median error over the 4 datasets. Mann-Whitney U-tests between the Bayesian results and the best comparative results confirmed statistical significance with $p < 0.05$ for Vol. J and Vol. K. The highest XMR-based MRI-to-physical registration and echo probe calibration errors were found for Vol. L, which might explain the poor performance of all motion estimation techniques.

Unrealistic motion estimates were reported for 26 out of 40 image-only estimates, while no unrealistic estimates were generated by the other estimation techniques. At least one unrealistic image-only motion estimate was reported for each dataset. An example of an unrealistic image-only estimate is shown in Figure 10. For Vol. J, one image-only registration resulted in a completely wrong estimate, with misalignment errors up to $100mm$.

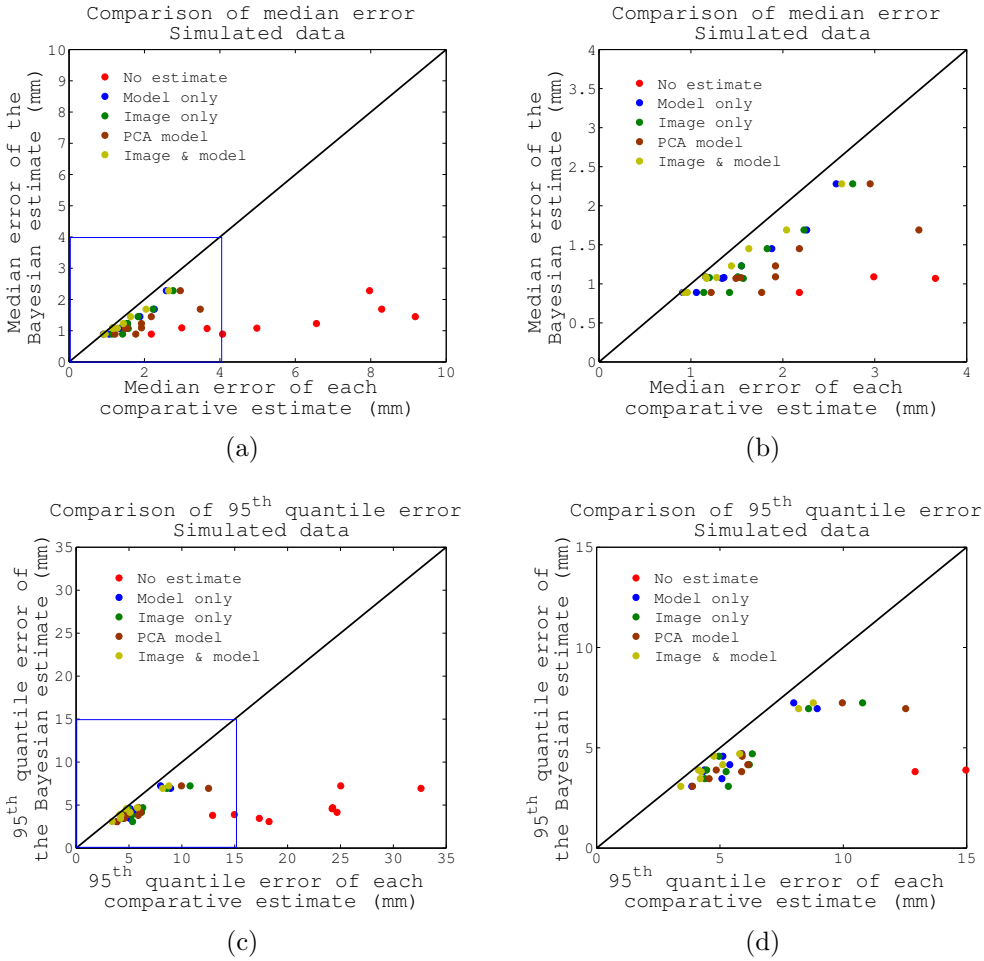


Figure 8: Accuracy error comparison between the Bayesian estimate (vertical axis) and each comparative estimate (horizontal axis) for simulated echo images. (a) Median error and (c) 95^{th} quantile error. Since there are some very high errors with no motion estimation, (b) and (d) show the error distributions in the blue boxes of Figure (a) and (c), respectively. The estimation error of the Bayesian technique was lower than any other comparative estimation technique.

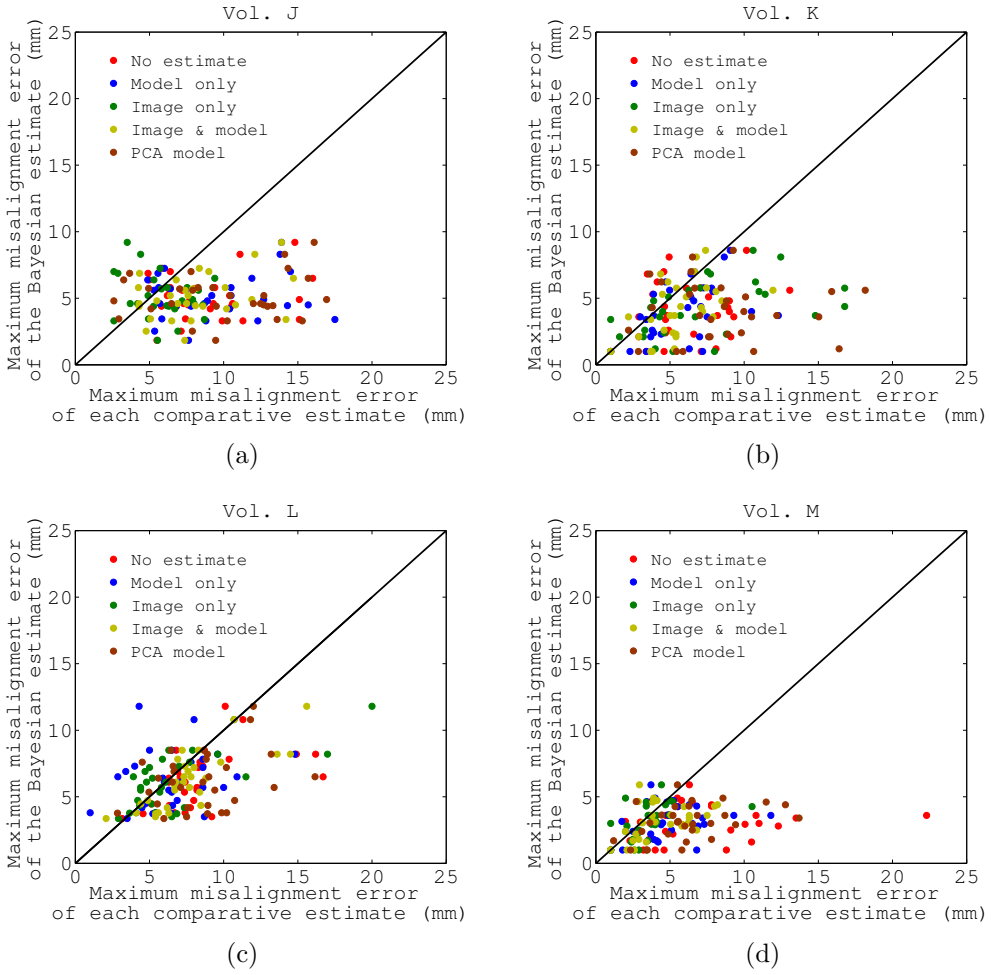
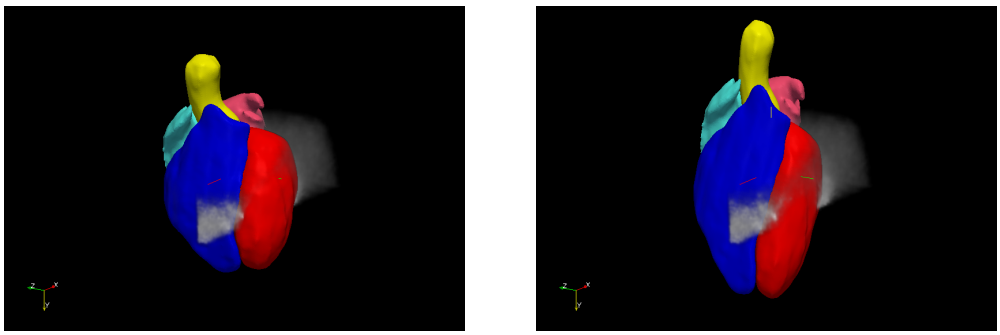


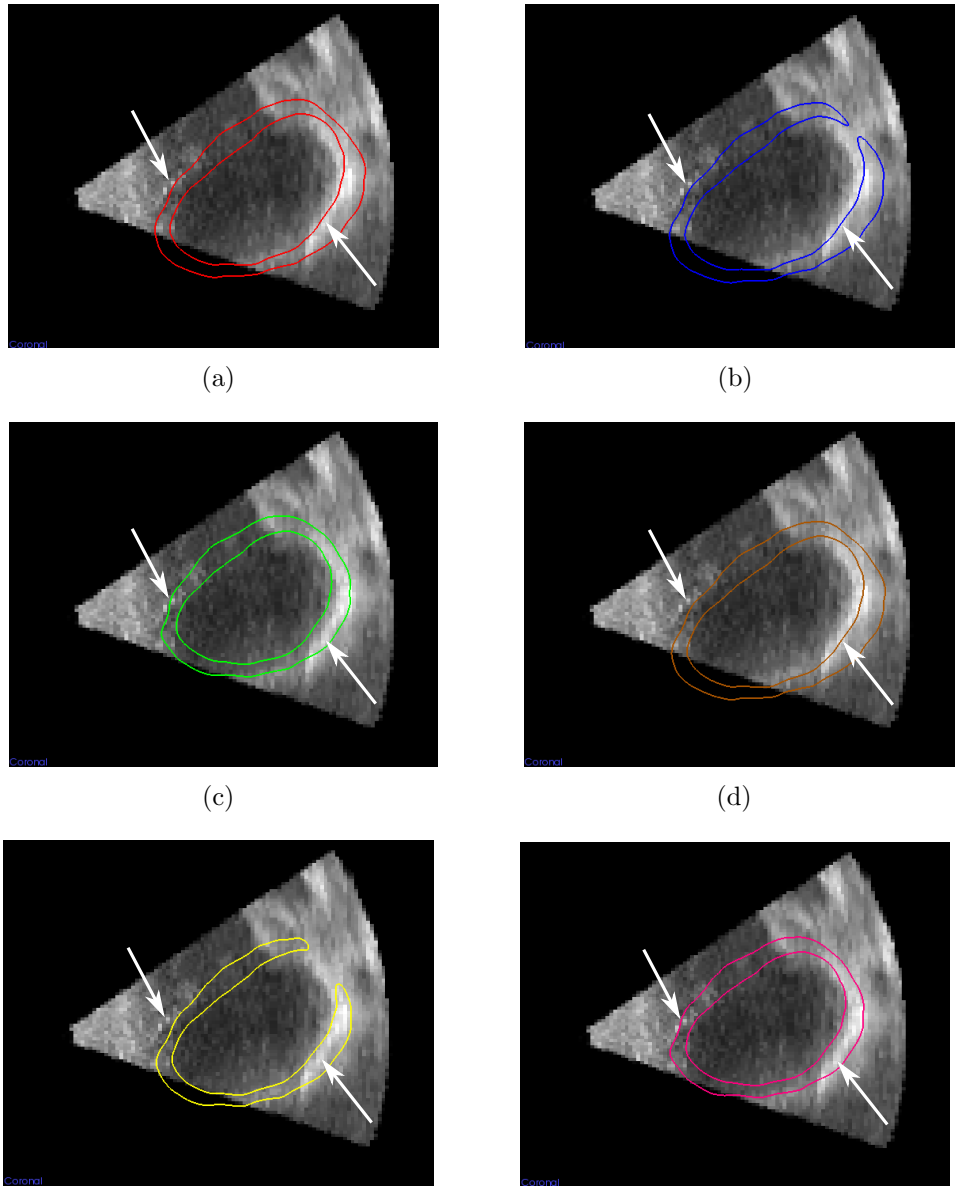
Figure 9: Maximum misalignment error comparison between the Bayesian estimate (vertical axis) and each comparative estimate (horizontal axis). Instead of median/95^{th} quantile, all misalignment errors are shown for (a) Vol. J, (b) Vol. K, (c) Vol. L and (d) Vol. M. For clarity of visualisation, three misalignments errors of the image-only estimation (48.8mm, 64mm and 100mm) for Vol. J (a) are not reported.

Figure 11 shows the comparison of the different motion estimates for a live 3D echo image. The segmentation of the LV myocardium is transformed using the different respiratory motion estimates and overlaid onto the corresponding live 3D echo image. A cross-section is used to visualise the misalignment error. For no motion estimate, model-only, PCA-based model and model & image estimation techniques residual misalignment errors can be seen at the posterior LV wall (bottom arrow) and inter-ventricular septum (top arrow). In the LV wall, the brightest voxels in the echo image represent the boundary between the LV myocardium and the left lung, therefore the external surface of the LV myocardium should be aligned with the boundary reflection. The lower misalignment errors are shown by the image-only 11(c) and Bayesian 11(f) estimation techniques. Figure 10(a) shows the surfaces of the four cardiac chambers and ascending aorta as segmented from the high resolution MRI while Figure 10(b) shows the cardiac surfaces as estimated by the image-only registration shown in Figure 11(c). The Bayesian motion estimation shows the best alignment while maintaining a realistic motion of the whole heart. Videos showing the performance of each estimation technique on a live 3D echo sequence are provided as supplementary material.



(a) (b)

Figure 10: (a) Segmentation of the heart as derived from the high resolution MRI image; (b) unrealistic image-only motion estimate.



(a) (b) (c) (d) (e) (f)

Figure 11: The motion estimates of the LV myocardium (coloured surfaces) are overlaid onto the corresponding live 3D echo image. Cross-sections are used to visualise the residual misalignment error. (a) No motion estimation, (b) model-only estimation, (c) image-only estimation, (d) PCA-based model, (e) model & image estimation and (f) Bayesian estimation are shown.

Subject	Median / 95^{th} quantile target error (mm)						Improv. (%)
	No estimate	*Model only*	*Image only*	*PCA model*	*Image & Model*	*Bayesian*	
Vol. A	2.99 / 24.26	1.16 / 5.89	1.51 / 6.31	1.92 / 5.86	1.16 / 5.80 †	1.09 / 4.70	6.0 / 18.9
Vol. B	3.66 / 12.91	1.34 / 4.32	1.57 / 5.25	1.49 / 5.88	1.17 / 4.24 †	1.07 / 3.81	8.5 / 10.1
Vol. C	4.07 / 18.25	1.06 / 3.86	1.14 / 5.34	1.22 / 3.89	0.96 / 3.41 †	0.89 / 3.08	7.3 / 9.7
Vol. D	4.98 / 14.97	1.36 / 4.37	1.20 / 4.46	1.54 / 4.85	1.28 / 4.13 †	1.08 / 3.89	15.6 / 5.8
Vol. E	7.97 / 25.03	2.58 / 7.99 †	2.76 / 10.78	2.95 / 9.96	2.64 / 8.78	2.28 / 7.24	11.6 / 9.4
Vol. F	8.29 / 32.62	2.26 / 8.94	2.23 / 8.59	3.48 / 12.53	2.04 / 8.19 †	1.69 / 6.95	17.2 / 15.1
Vol. G	9.18 / 24.68	1.88 / 5.40	1.83 / 6.19	2.18 / 6.13	1.63 / 5.11†	1.45 / 4.16	11.0 / 18.6
Vol. H	6.56 / 24.21	1.55 / 5.12	1.55 / 4.96	1.92 / 5.90	1.44 / 4.75 †	1.23 / 4.58	14.6 / 3.6
Vol. I	2.18 / 17.33	0.91 / 5.08	1.42 / 4.39	1.77 / 4.56	0.92 / 4.20 †	0.89 / 3.46	3.4 / 17.6

Table 2: Evaluation results of Bayesian estimation using simulated echo images. Median and 95^{th} quantile values of target error are reported. The last column shows the percentage improvement in median and 95^{th} quantile target error of the Bayesian technique over the best comparative estimate, denoted by a †.

Subject	Median / 95^{th} quantile maximum error in misalignment (mm)						Improv. (%)
	No estimate	*Model-only*	*Image-only*	*PCA model*	*Image & Model*	*Bayesian*	
Vol. J	7.65 / 15.10	8.75 / 15.70	6.19 / 64.00	9.87 / 15.92	7.70 / 14.20 †	4.85 / 8.30	37.0 / 41.5
Vol. K	5.52 / 10.16	6.35 / 10.50	5.98 / 16.78	8.09 / 16.17	5.35 / 8.50 †	3.85 / 8.10	28.0 / 4.7
Vol. L	7.87 / 16.20	6.48 / 10.90 †	5.74 / 17.00	8.75 / 13.31	7.31 / 14.50	6.10 / 10.80	5.9 / 0.9
Vol. M	5.66 / 13.5	4.65 / 9.30	3.51 / 5.19 †	5.90 / 12.19	3.87 / 8.00	3.07 / 4.90	12.5 / 5.6

Table 3: Evaluation results of Bayesian estimation using real echo images. Median and 95^{th} quantile values of the maximum misalignment error are reported. The last column shows the percentage improvement in median and 95^{th} quantile target error of the Bayesian technique over the best comparative estimate, denoted by a †.

5. Discussion and conclusions

We have presented a novel Bayesian model for respiratory motion estimation and compensation in image-guided cardiac interventions. Our method combines a MRI-derived affine motion model of the heart with intra-procedure live 3D echocardiography in a Bayesian approach. The novelties of the proposed method lie in its probabilistic formulation, that embeds a measure of confidence in its estimates and resolves the uncertainty related to the variability of respiratory motion and the potentially conflicting information carried by the model and the imaging data. Moreover, the use of probability density functions overcomes the limitations of motion models proposed so far, allowing the Bayesian motion estimate to deviate from the predetermined model only estimate. By determining the optimal value of the only free parameter σ_l^2, the Bayesian technique adapts automatically to the quality of the echo images. The Bayesian method was initially proposed in Peressutti et al. (2012); this paper presented evaluation on more 3D simulated echo datasets and 3D real echo datasets, a heteroscedastic variance function for the prior probability and an automatic method for choosing σ_l^2.

Simulated echo images with realistic MRI-derived motion fields were employed to perform a thorough accuracy/robustness evaluation with known, gold-standard, motion fields. Results showed mean/max accuracy improvements in motion estimation of 10.6%/18.9% over the best comparative estimation technique. Although echo image artefacts such as spatial distortion and incidence angle were not included in the simulation process, realism of the simulated echo images was improved by including real speckle textures and Gaussian noise. Even though SNR_d is a simple measure of echo image quality, the comparable SNR_d values achieved for simulated and real echo images suggest that similar image quality was achieved. The lack of the rib cage simulation was compensated by the selection of an echo beam positioning that resembled realistic acoustic windows.

Real echo imaging data enabled evaluation on data similar to those which would be acquired in a clinical environment. For the 4 volunteer datasets considered, results showed the Bayesian method to outperform all comparative approaches, producing the lowest misalignment errors with mean/max improvements in estimation accuracy of 20.8%/41.5%. Due to more sources of inaccuracy, such as errors in the probe and XMR suite calibration (Rhode et al., 2003, 2005; Ma et al., 2009), synchronisation, evaluation assessment and errors due to echo image artefacts not included in the simulation, errors

were higher for real echo images compared to simulated echo. In addition, the respiratory bellows signal was subject to amplitude drifts, which directly affect model-only estimates. Spirometry has also been proposed in the literature as a surrogate signal but is impractical to use in most cardiac interventions, causing high discomfort to the patient. Echo images employed in this study were acquired in a supine position, where the position of the heart is not ideal to be imaged. Better quality echo images could further improve the Bayesian motion estimates. Future work will investigate methods to identify anatomical regions that maximise the information gained from the echo images about the variability of respiratory motion.

This paper has also provided a comparative study on motion estimation techniques available in the literature. Results from both simulated and real echo images have confirmed the findings presented in King et al. (2010), showing the combination of a motion model and echo data to be more robust than the echo image-only estimation, which, despite being accurate in the area of coverage of the echo images, is not reliable for a whole heart motion estimation. Furthermore, echo image-only registration was unstable if a suitably close starting estimate was not provided. The improvement gained by the Bayesian method over the model and image technique represents the amount of respiratory variability that Bayesian estimates are able to capture. Since this variability is highly subject dependent, so is the level of improvement. Therefore, results were particularly improved for individuals where breathing patterns altered significantly. This adaptability to changes in breathing patterns means that our method has the potential to widen the patient population who could benefit from image guidance. Concerning model only estimates, $\tilde{\boldsymbol{\theta}}(s)$, inaccuracies affecting the respiratory surrogate measurement in building and applying the model directly affected the estimation accuracy, resulting in higher errors, particularly for the real echo experiments. The performance of the PCA-based motion model was comparable to the image-only estimation, even though no initial estimate was provided to the PCA-based model. The lack of the surrogate information along with artefacts and a low quality of the echo images may be the cause of the poor performance of the PCA-based motion model. Accuracy errors corresponding to no motion estimation highlight the need of motion compensation techniques to fulfil clinical requirements (Linte et al., 2010).

A simple but effective formulation for the automatic determination of the optimal weighting of the likelihood and prior terms was also proposed. The linear relationship between σ_l^2 and SNR_d simplifies the application of

the Bayesian model, which only requires the knowledge of the SNR_d of the echo image acquired. The formulation for the optimal σ_l^2 is meant to be valid for any new patient dataset and for this reason we used a leave-one-out technique to derive the coefficients a_1 and a_2 for the datasets considered. However, the typical values of a_1 and a_2 need to be evaluated on a larger set of volunteers/patients to account for variations affecting the echo imaging acquisition. In this study, the SNR_d was derived from the echo-only image registration and, given the increasing availability of 3D live-stream echo images and parallelisation algorithms, the SNR_d could be rapidly computed and made available for the Bayesian estimation. The proposed method for determining the optimal value of σ_l^2 used a segmentation of the LV cavity and myocardium derived from the high resolution MRI image, which were then warped to each live echo image using the MRI to patient rigid registration and the affine echo-only registration. The MRI to patient rigid registration is available from the XMR suite and the associated rigid registration errors are estimated to be approximately 2 mm (Rhode et al., 2005). The affine echo-only image registration was used to warp the LV segmentations to the respiratory position of each live 3D echo image since, as previously shown by King et al. (2010) and as confirmed by our results, echo-only registrations produce reasonably accurate estimations within the echo FOV (although much less so outside the FOV). For these reasons, the proposed approach has proved to be accurate enough to make an approximate estimate of SNR_d. However, alternative methods that rely on echo images only, such as histogram clustering techniques, could be employed instead (Noble and Boukerroui, 2006).

As already mentioned in Section 2.1.1, the assumption of statistical independence of the affine parameters is a simplification and may not be strictly valid. However, accounting for the unknown dependencies between parameters would require a large amount of data, increasing the complexity of the clinical acquisition protocol and the Bayesian estimation. Therefore, the proposed formulation represents a simplified but effective description of the prior information about subject-specific respiratory motion. The time required to build the prior probability term depends on the affine registration algorithm used. In our case, from the acquisition of the dynamic MRI images, the prior probability term was built in less than 20 minutes. This time lag is shorter than the time required to move the patient from the MRI scanner to the operative X-ray table and to prepare them for the intervention. The Bayesian method was implemented in C++ with an average CPU execution time of 120 seconds on an 8-core i7-2600 processor per echo image. Therefore the

technique is not currently real-time. However, code optimisation was not the focus of this paper and will need to be considered for real-time application.

Echo imaging is not currently used routinely in image-guided cardiac interventions, but there is an increasing trend toward incorporating echo information in the procedural workflow (Linte et al., 2008; Wein et al., 2009; Gao et al., 2012). Our method provides a framework for using echo data to estimate respiratory motion, which could be used in combination with a robotic arm (Ma et al., 2010) for automatic acquisition of the 3D live echo images during the procedure. In principle, the Bayesian model could be used without the measurement of the respiratory surrogate during the procedure, since the surrogate is optimised as part of the maximisation of the posterior probability. This would reduce the impact on the clinical workflow.

Future work will evaluate the Bayesian model on patient data and transoesophageal echo images. In terms of clinical applications, the Bayesian model has the potential to improve guidance in many cardiac interventions, such as catheterisation procedures for EP studies, valve implantation, or cardiac stem cell implantation, where the accuracy requirement is very high. Finally, this work showed results for cardiac respiratory motion, but could be applied to improve the accuracy of image-guided interventions applied to any other organ affected by respiratory motion.

Download

The MRI and echo datasets of Vol. J, Vol. K, Vol. L, and Vol. M and the description of how to use them are available to download at `https://www.isd.kcl.ac.uk/BayesianModel/`.

Acknowledgements

This work was funded by EPSRC grant EP/H046410/1. Erik-Jan Rijkhorst undertook this research whilst funded by EPSRC grant EP/F025750/1. Dean Barratt is funded by a Royal Academy of Engineering/EPSRC Research Fellowship. The authors acknowledge financial support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust.

Baun, J., 2010. Principles of General & Vascular Sonography. ProSono.

Blackall, J.M., Ahmad, S., Miquel, M.E., McClelland, J.R., Landau, D.B., Hawkes, D.J., 2006. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. Physics in Medicine and Biology 53, 4147–4169.

Blackall, J.M., Penney, G.P., King, A.P., Hawkes, D.J., 2005. Alignment of sparse freehand 3-D ultrasound with preoperative images of the liver using models of respiratory motion and deformation. IEEE Transactions on Medical Imaging 24, 1405–1416.

Blanc, R., Syrkina, E., Szekely, G., 2009. Estimating the confidence of statistical model based shape prediction. Proc of IPMI 21, 602–613.

Buerger, C., Schaeffter, T., King, A.P., 2011. Hierarchical adaptive local affine registration for fast and robust respiratory motion estimation. Medical Image Analysis 15, 551–564.

Cleary, K., Peters, T.M., 2010. Image-guided interventions: Technology review and clinical applications. Annual Review of Biomedical Engineering 12, 119–142.

Davidian, M., Carroll, R.J., 1987. Variance function estimation. Journal of the American Statistical Association 82, 1079–1091.

De Buck, S., Maes, F., Ector, J., Bogaert, J., Dymarkowski, S., Heidbuchel, H., Suetens, P., 2005. An augmented reality system for patient-specific guidance of cardiac catheter ablation procedures. Medical Imaging, IEEE Transactions on 24, 1512 –1524.

Ector, J., Buck, S.D., Huybrechts, W., Nuyens, D., Dymarkowski, S., Bogaert, J., Maes, F., Heidbuechel, H., 2008. Biplane three-dimensional augmented fluoroscopy as single navigation tool for ablation of atrial fibrillation: Accuracy and clinical value. Heart Rhythm 5, 957 – 964.

Efros, A.A., Freeman, W.T., 2001. Image quilting for texture synthesis and transfer, in: ACM (Ed.), SIGGRAPH, Association for Computing Machinery., pp. 341–346.

Gao, G., Penney, G., Ma, Y., Gogin, N., Cathier, P., Arujuna, A., Mortona, G., Caulfield, D., Gilld, J., Rinaldi, C.A., Hancock, J., Redwood, S., Thomas, M., Razavi, R., Gijsbers, G., Rhode, K., 2012. Registration of 3D

trans-esophageal echocardiography to X-ray fluoroscopy using image-based probe tracking. Medical Image Analysis 16, 38–49.

Geman, S., McClure, D.E., 1985. Bayesian image analysis: An application to single photon emission tomography., in: 1985 Proceedings of the American Statistical Association. Statistical Computing Section, pp. 12–18.

Grau, V., Becher, H., Noble, J.A., 2007. Registration of multiview real-time 3-D echocardiographic sequences. IEEE Transactions on Medical Imaging 26, 1154–1165.

Hanson, K.M., 1993. Introduction to Bayesian image analysis, in: SPIE, Medical Imaging : Image Processing, pp. 716–731.

Hawkes, D., Barratt, D., Blackall, J., Chan, C., Edwards, P., Rhode, K., Penney, G., McClelland, J., Hill, D., 2005. Tissue deformation and shape models in image-guided interventions: a discussion paper. Medical Image Analysis 9, 163 – 175.

Keall, P.J., Mageras, G.S., Balter, J.M., Emery, R.S., Forster, K.M., Jiang, S.B., Kapatoes, J.M., Low, D.A., Murphy, M.J., Murray, B.R., Ramsey, C.R., Van Herk, M.B., Vedam, S.S., Wong, J.W., Yorke, E., 2006. The management of respiratory motion in radiation oncology, report of AAPM Task Group 76. Medical Physics 33, 3874+.

King, A., Boubertakh, R., Rhode, K., Ma, Y., Chinchapatnam, P., Gao, G., Tangcharoen, T., Ginks, M., Cooklin, M., Gill, J., Hawkes, D., Razavi, R., Schaeffter, T., 2009. A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures. Medical Image Analysis 13, 419–431.

King, A., Buerger, C., Tsoumpas, C., Marsden, P., Schaeffter, T., 2012. Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator. Medical Image Analysis 16, 252 – 264.

King, A., Jansen, C., Rhode, K., Caulfield, D., Razavi, R., Penney, G., 2010. Respiratory motion correction for image-guided cardiac interventions using 3-D echocardiography. Medical Image Analysis 14, 21–29.

Klinder, T., Lorenz, C., Ostermann, J., 2010. Prediction framework for statistical respiratory motion modeling, in: Medical Image Computing and Computer-Assisted Intervention ? MICCAI 2010.

Knecht, S., Skali, H., ONeill, M.D., Wright, M., Matsuo, S., Chaudhry, G.M., Haffajee, C.I., Nault, I., Gijsbers, G.H.M., Sacher, F., 2008. Computed tomography-fluoroscopy overlay evaluation during catheter ablation of left atrial arrhythmia. Europace European journal of pacing arrhythmias and cardiac electrophysiology 10, 931–938.

Kondermann, C., Mester, R., Garbe, C., 2008. A statistical confidence measure for optical flows, in: Proceedings of the 10th European Conference on Computer Vision: Part III, Springer-Verlag. pp. 290–301.

Kremkau, F.W., Taylor, K.J., 1986. Artifacts in ultrasound imaging. Journal of Ultrasound in Medicine 5, 227–237.

Larrabide, I., Omedas, P., Martelli, Y., Planes, X., Nieber, M., Moya, J., Butakoff, C., Sebastian, R., Camara, O., Craene, M.D., Bijnens, B., Frangi, A., 2009. GIMIAS: an open source framework for effcient development of research tools and clinical prototypes. Functional Imaging and Modeling of the Heart, ser. Lecture Notes in Computer Science 5528, 417–426.

Linte, C., Moore, J., Peters, T., 2010. How accurate is accurate enough? A brief overview on accuracy considerations in image-guided cardiac interventions, in: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, pp. 2313 –2316.

Linte, C., Moore, J., Wiles, A., Wedlake, C., Peters, T., 2008. Virtual reality-enhanced ultrasound guidance: a novel technique for intracardiac interventions. Computer Aided Surgery 13, 82–94.

Ma, Y., Penney, G.P., Bos, D., Frissen, P., Rinaldi, C.A., Razavi, R., Rhode, K.S., 2010. Hybrid echo and X-ray image guidance for cardiac catheterization procedures by using a robotic arm: a feasibility study. Physics in Medicine and Biology 55, N371.

Ma, Y.L., Penney, G.P., Rinaldi, C.A., Cooklin, M., Razavi, R., Rhode, K.S., 2009. Echocardiography to magnetic resonance image registration for use in image-guided cardiac catheterization procedures. Physics in Medicine and Biology 54, 5039?5055.

Manke, D., Nehrke, K., Bornert, P., 2003. Novel prospective respiratory motion correction approach for free-breathing coronary MR angiography using a patient-adapted affine motion model. Magnetic Resonance in Medicine 50, 122–131.

Noble, A.J., Boukerroui, D., 2006. Ultrasound image segmentation : A survey. IEEE Transactions on Medical Imaging 25, 987–1010.

Noble, A.J., Navab, N., Becher, H., 2011. Ultrasonic image analysis and image-guided interventions. Interface Focus 1, 673–685.

Peressutti, D., Rijkhorst, E.J., Barratt, D.C., Penney, G.P., King, A.P., 2012. Estimating and resolving uncertainty in cardiac respiratory motion modelling, in: Biomedical Imaging: From Nano to Macro, 2012 IEEE International Symposium on, pp. 262–265.

Perrin, D.P., Vasilyev, N.V., Novotny, P., Stoll, J., Howe, R.D., Dupont, P.E., Salgo, I.S., del Nido, P.J., 2009. Image guided surgical interventions. Current Problems in Surgery 46, 730 – 766.

Rhode, K., Hill, D., Edwards, P., Hipwell, J., Rueckert, D., Sanchez-Ortiz, G., Hegde, S., Rahunathan, V., Razavi, R., 2003. Registration and tracking to integrate X-ray and MR images in an XMR facility. Medical Imaging, IEEE Transactions on 22, 1369 –1378.

Rhode, K.S., Sermesant, M., Brogan, D., Hegde, S., Hipwell, J., Lambiase, P., Rosenthal, E., Bucknall, C., Qureshi, S.A., Gill, J.S., Razavi, R., Hill, D.L.G., 2005. A system for real-time XMR guided cardiovascular intervention. IEEE Transactions on Medical Imaging 24, 1428–1440.

Rijkhorst, E.J., Heanes, D., Odille, F., Hawkes, D., Barratt, D., 2010. Simulating dynamic ultrasound using mr-derived motion models to assess respiratory synchronisation for image-guided liver interventions, in: IPCAI'10 Proceedings of the First international conference on Information processing in computer-assisted interventions.

Savill, F., Schaeffter, T., King, A., 2011. Assessment of input signal positioning for cardiac respiratory motion models during different breathing patterns, in: Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, pp. 1698 –1701.

Schneider, M., Sundar, H., Liao, R., Hornegger, J., Xu, C., 2010. Model-based respiratory motion compensation for image-guided cardiac interventions, in: Computer Vision and Pattern Recognition, 2010 IEEE Conference on.

Seber, G.A.F., Lee, A.J., 2003. Linear Regression Analysis, 2nd edition. Wiley.

Shechter, G., Shechter, B., Resar, J.R., Beyar, R., 2005. Prospective motion correction of X-ray images for coronary interventions. IEEE Transactions on Medical Imaging 24, 441–450.

Sra, J., Narayan, G., Krum, D., Malloy, A., Cooley, R., Bhatia, A., Dhala, A., Blanck, Z., Nangia, V., Akhtar, M., 2007. Computed tomography-fluoroscopy image integration-guided catheter ablation of atrial fibrillation. Journal of Cardiovascular Electrophysiology 18, 409–414.

Taron, M., Paragios, N., Jolly, M.P., 2005. Uncertainty-driven non-parametric knowledge-based segmentation : the corpus callosum case, in: IEEE Workshop on Variational and Level Set Methods in Computer Vision, pp. 1–24.

Thijssen, J.M., 2003. Ultrasonic speckle formation, analysis and processing applied to tissue characterization. Pattern Recognition Letters 24, 659–675.

Wein, W., Camus, E., John, M., Diallo, M., Duong, C., Al-Ahmad, A., Fahrig, R., Khamene, A., Xu, C., 2009. Towards guidance of electrophysiological procedures with real-time 3D intracardiac echocardiography fusion to C-arm CT, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2009.

Yu, H., Fahrig, R., Pelc, N., 2005. Co-registration of X-ray and MR fields of view in a hybrid XMR system. Journal of Magnetic Resonance Imaging 22, 291–301.